

Automatic coding and keyboard indexing method for Chinese database

Publication number:CN1258052

Publication date:2000-06-28

Inventor:MAI GUANGSHU (CN)

Applicant:MAI GUANGSHU (CN)

Classification: - international:**G06F17/30; G06F17/30**; (IPC1-7): G06F17/30 - European:

Application number:CN19981022250 19981221

Priority number(s):CN19981022250 19981221

Abstract of **CN1258052**

An automatic coding and keyboard indexing method for Chinese database features that the Chinese characters are decomposed to several basic strokes,when literal document is input the electronic information processor can automatically take the strokes at particular sequential positions and use them as the codes of the data terms, and when indexing, the codes of the data terms are input and the electronic information processor can compare them with the codes in database and call out the data terms according to their matches. It is suitable for mobile telephone and pager with key pad.

Data supplied from the **esp@cenet** database - Worldwide

[12] 发明专利申请公开说明书

[21] 申请号 98122250.1

[43] 公开日 2000 年 6 月 28 日

[11] 公开号 CN 1258052A

[22] 申请日 1998.12.21 [21] 申请号 98122250.1

[71] 申请人 麦广树

地址 528244 广东省南海市里水镇河村江边

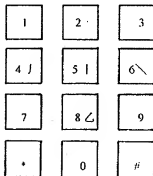
[72] 发明人 麦广树

权利要求书 7 页 说明书 14 页 附图页数 4 页

[54] 发明名称 中文数据库的自动编码及键盘检索方法

[57] 摘要

中文数据库的自动编码及键盘检索方法,其特征在于将汉字分解为若干基本笔画,在输入文字资料的过程中,电子信息处理器自动取出若干个汉字特定序位上的笔画作为该数据项的编码,检索时输入该数据项的编码,电子信息处理器根据所输入的编码与数据库中各数据项编码的各码位比较后,将与所输入编码具有相同码位数目的多少,先后调出各数据项。本发明可与汉字输入方法无关,适合于移动电话、传呼机等小键盘产品上作为数据库资料的检索。



ISSN 1008-4274

权 利 要 求 书

1、一种以汉字字符串作检索入口的数据库自动编码及键盘检索方法，其特征在于：

A、对汉字的笔画进行分类；

B、将分类得到的汉字基本笔画定义在键盘相应的键位上；

C、根据依A分类所得的汉字基本笔画及汉字的笔画书写顺序取出汉字特定序位上的若干笔画，取出的笔画即定义为该汉字的编码，该汉字编码、区位码和汉字储存在电子信息处理器的存储器中；

D、电子信息处理器根据C的汉字编码规则取出作为数据库检索入口而输入到该电子信息自理器中的汉字字符串中的若干个汉字的编码，并将所取得的汉字的编码组合形成相应数据项的检索码；

E、检索时，根据B的汉字基本笔画在键盘位上的定义从键盘输入需检索的数据项的依D所定义的检索码，电子信息处理器将键盘输入的检索码的各个码位分别与数据库中各数据项的检索码的对应码位比较后，调出具有最多相同码位的数据项；用移项功能键按比较后码位相同的数目多少，先多后少地顺序调出各数据项。

2、一种以汉字字符串作检索入口的数据库自动编码及键盘检索方法，其特征在于：

A、对汉字的笔画进行分类；

B、将分类得到的汉字基本笔画定义在键盘相应的键位上;

C、根据依A分类所得的汉字基本笔画及汉字的笔画书写顺序取出汉字特定序位上的若干笔画, 取出的笔画即定义为该汉字的编码, 该汉字的编码隐藏在汉字输入码中, 汉字输入码和汉字储存在电子信息处理器的存储器中;

D、电子信息处理器根据C的汉字编码规则取出作为数据库检索入口而输入到该电子信息处理器中的汉字字串中的若干个汉字的编码, 并将所取得的汉字的编码组合形成相应数据项的检索码;

E、检索时, 根据B的汉字基本笔画在键盘位上的定义从键盘输入需检索的数据项的依D所定义的检索码, 电子信息处理器将键盘输入的检索码的各个码位分别与数据库中各数据项的检索码的对应码位比较后, 调出具有最多相同码位的数据项; 用移项功能键按比较后码位相同的数目多少, 先多后少地顺序调出各数据项。

3、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法, 其特征在于汉字的基本笔画可分为“横”, “竖”, “撇”, “捺”; 分别用以下符号表示: “一” “|”, “丿”, “\”; 其中“点”归在“捺”内, 所有“折”笔根据其第一折弯的前段分别归入“一”, “|”, “丿”, “\”。

4、根据权利要求1或2所述的以汉字字串作检索入口的

数据库自动编码及键盘检索方法，其特征在于汉字的基本笔画可分为“横”，“竖”，“撇”，“点”，“捺”；分别用以下符号表示，“一”，“丨”，“丿”，“丶”，“㇏”，所有“折”笔根据其第一折弯的前段分别归入“一”，“丨”，“丿”，“丶”，“㇏”。

5、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于汉字的基本笔画可分为“横”，“竖”，“撇”，“捺”，“折”；分别用以下符号表示，“一”，“丨”，“丿”，“丶”，“乙”；“捺”归在“点”内。

6、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于汉字的基本笔画可分为“横”，“竖”，“撇”，“点”，“捺”，“折”；分别用以下符号表示，“一”，“丨”，“丿”，“丶”，“㇏”，“乙”。

7、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于汉字的基本笔画可分为“横”，“竖”，“撇”，“捺”，“顺折”，“反折”；分别用以下符：“一”，“丨”，“丿”，“丶”，“㇏”，“㇔”，“㇕”；其中“点”归在“捺”内，收笔时折向为顺时针方向的为“顺折”，反之为“反折”。

8、根据要得要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于汉字的基本笔画

可分为“横”，“竖”，“撇”，“点”，“捺”，“顺折”，“反折”；分别用以下符合表示：“一”，“|”，“丿”，“丶”，“\”，“┐”，“└”；其中收笔时折向为顺时针方向的为“顺折”，反之为“反折”。

9、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于汉字的基本笔画可分为“横”，“竖”，“撇”，“捺”，“单折”，“多折”；分别用以下符号表示，“一”，“|”，“丿”，“丶”，“┐”，“乙”；其中“点”归在“捺内”，“单折”为只有一个弯的折笔，“多折”为具有二个或二个以上转弯的折笔。

10、根据权利要求1或2所述的以汉字字串作检索入口的数据自动编码及键盘检索方法，其特征在于汉字的基本笔画可分为“横”，“竖”，“撇”，“点”，“捺”，“单折”，“多折”；分别用以下符号表示，“一”，“|”，“丿”，“丶”，“\”，“┐”，“乙”其中“单折”为只有一个转弯的折笔，“多折”为具有二个或二个以上转弯的折笔。

11、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于经分类所得的汉字的每个基本笔画用单个按键定义。

12、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于经分类所得的汉字的每个基本笔画用二键次或二键次以上的按键组结合笔画的

形态来定义。

13、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于对汉字进行编码时同时取出汉字的首笔笔画和末笔笔画，首笔笔画在前，末笔笔画在后，两者组合在一起定义为该汉字的编码；对只有一个笔画的汉字只取该笔画为该汉字的编码。

14、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于对汉字进行编码时只取出汉字的首笔笔画，首笔笔画即定义为该汉字的编码。

15、据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于对汉字进行编码时取出汉字的末笔笔画，末笔笔画即定义为该汉字的编码。

16、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于对汉字进行编码时同取出汉字的首笔笔画和第二笔笔画，首笔笔画在前，第二笔笔画在后，两者结合在一起定义为该汉字的编码；对只有一个笔画的汉字只取该笔画为该汉字的编码。

17、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于电子信息处理器取出作为数据库检索入口而输入到该电子信息处理器的汉字字串中的每个汉字的编码，取出的汉字编码根据相应汉字在汉字字串中的顺序组合形成相应数据项的检索码。

18、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于电子信息处理器也可以取出作为数据库检索入口而输入到该电子信息处理器的汉字字串中的首位汉字和末位汉字的编码，取出的汉字根据相应汉字在汉字字串中的顺序组合形成相应数据的检索码。

19、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于电子信息处理器取出作为数据库检索入口而输入到该电子信息处理的汉字字串中的前两个汉字的编码，对只有一个汉字的字串只取该汉字的编码，取出的汉字编码根据相应汉字在汉字字串中的顺序组合形成相应数据项的检索码。

20、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于电子信息处理器取出作为数据库检索入口而输入到该电子信息处理器的汉字字串中的前三个汉字的编码，对只有三个或三个以下汉字字串的，取全部汉字的编码，取出的汉字编码根据相应汉字在汉字字串中的顺序组合形成相应数据项的检索码。

21、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于电子信息处理器取出作为数据库检索入口而输入到该电子信息处理器的汉字字串中的前四个汉字的编码，对只有四个或四个以下汉字字串的，取全部汉字的编码，取出的汉字编码根据相应汉字在汉字字串

中的顺序组合形成相应数据项的检索码。

22、根据权利要求1或2所述的以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征就在于所述汉字包括汉字基本笔画：“一”，“丨”，“丿”，“丶”，“㇏”，“㇚”，“㇚”
“乙”和汉字部首：“𠂇”，“木”，“彳”，“扌”，“口”，“火”，“土”，“禾”，“竹”，“文”，“灬”等。

中文数据库的自动编码及键盘检索方法

本发明涉及一种以汉字字串作检索入口的数据库自动编码及键盘检索方法。

计算机汉字编码方法已经很多，但是专门针对以检索资料为主的汉字编码方法却很少。目前，利用电子计算机等电子信息处理器检索以汉字字串作为检索入口的数据库资料时，一般需要输入该汉字字串作为检索索引。或者只能显示所储存的数据库资料，然后人工选择所需的数据项。无论前者或后者都要操作者花费较大的精力，对不熟悉汉字输入方法的操作者而言，困难将会更大。上述缺陷使得以汉字字串作检索入口的数据库技术的应用受到了极大的限制，目前，手提电话、传呼机、电子记事簿、个人数字助理器、个人移动信息处理器等产品已经可以通过无线电或红外线与个人电脑进行资料传输，解决了这些小键盘产品的大量汉字资料输入的速度问题。而它们的主要作用——资料检索，因键盘太少只能停留在菜单浏览选择或直接输入资料序号的方法上，其缺点是菜单浏览选择检索方法速度太慢，而直接输入资料序号则要死记大量的与资料对座的序号，十分不方便。即使用五笔画输入法输入汉字来检索，因输入每个汉字的重码率太高，速度同样很慢。当前，随着信息电子技术的发展，个人电脑、手提电话、电子记事簿、传呼机等产品的应用已较为普及，而这些产品的键盘都不太大，并且是以检索资料为主，所以如何有效地扩展这些产

品的功能，使其同时可以方便、快捷地检索储存于其中的汉字信息是提高汉字信息技术、增强产品市场竞争力的重要手段。

鉴于汉字信息处理技术应在提手电话、传呼机、个人移动信息处理器、个人数字助理器、电子记事簿等小键盘产品上所遇到的问题，本人发明了一种数据库的自动编码及键盘检索方法。

本发明的目的在于给小键盘电子信息处理产品提供一种以汉字字符串作检索入口的数据库自动编码及其方便、快捷的键盘容错检索方法。使人们在检索资料时只需直接输入经电子信息处理器自动编码的数据项资料的检索码，检索具有容错功能，使用户在输错检索码时无需重新输码也可以用移项功能键快速找到所需的数据项。它与所采用的汉字输入方法无关。给可以通过无线电、红外线等方法输入汉字的传呼机、手提电话、个人移动信息处理器、电子记事簿、电话等产品提供一种键数少、键盘简单、易于操作，便于记忆，操作速度快捷的汉字数据库自动编码及键盘检索方法。避免了传呼机、手提电话、个人移动信息处理器等产品用输入汉字字符串的方法检索资料时因输入汉字而遇到的键数少而重码率高，重码率低则键数多而且使用复杂的矛盾。克服了上述产品用数据项序号检索资料时，数据项序号与相应资料内容关系难以记忆而用菜单浏览方法检索资料时检索速度又太慢的缺点。

为达到上述目的，本发明的第一种技术方案是这样的：以汉字字符串作检索入口的数据库自动编码及键盘检索方法，其特征在于：

- A、对汉字的笔画进行分类；
- B、将分类得到的汉字基本笔画定义在键盘相应的键位上；

C、根据依A分类所得的汉字基本笔画及汉字的笔画书写顺序取出汉字特定序位上若干笔画，取出的笔画即定义为该汉字的编码，该汉字编码、区位码和汉字储存在电子信息处理器的存贮器中；

D、电子信息处理器根据C的汉字编码规则取出作为数据库检索入口而输入到该电子信息处理器中的汉字字串中的若干个汉字的编码，并将所取得的汉字的编码组合形成相应数据项的检索码；

E、检索时，根据B有的汉字基本笔画在键盘键位上的定义从键盘输入需检索的数据项的依D所定义的检索码，电子信息处理器将键盘输入的检索码的各个码位分别与数据库中各数据项的检索码的对应码位比较后，调出具有最多相同码位的数据项；用移项功能键按比较后码位相同的数目多少，先多后少地顺序调出各数据项。

本发明的第二种技术方案是这样的：以汉字字串作检索入口的数据库自动编码及键盘检索方法，其特征在于：

A. 对汉字的笔画进行分类；

B. 将分类得到的汉字基本笔画定义在键盘相应的键位上；

C. 根据依A分类所得的汉字基本笔画及汉字的笔画书写顺序取出汉字特字序位上的若干笔画，取出的笔画即定义为该汉字的编码，该汉字的编码隐藏在汉字输入码中，汉字输入码和汉字储存在电子信息处理器的存贮器中；

D. 电子信息处理器根据C的汉字编码规则取出作为数据库检索入口而输入到电子信息处理器中的汉字字串中的若干个汉字的编码，并将所取得的汉字的编码组合形成相应数据项的检索码；

E. 检索时, 根据 B 的汉字基本笔画在键盘键位上的定义从键盘输入需检索的数据项的依 D 所定义的检索码, 电子信息处理器将键盘输入的检索码的各个码位分别与数据库中各数据项的检索码的对应码位比较后, 调出具有最多相同码位的数据项; 用移项功能键按比较后码位相同的数目多少, 先多后少地顺序调出各数据项。

按照上述技术方案, 汉字笔画的分类基本是依据汉字的书写及认知习惯进行的。汉字基本笔画定义在键盘相应的键位上, 该键位可以用相应汉字基本笔画进行标示, 也可以和字、字母或其他符号进行标示。上述技术方案中所称的汉字特定序位上的笔画是指根据汉字的笔画书写顺序确定出的该汉字的第 1 笔(首笔)、第 2 笔、……第 n 笔或者该汉字的末笔。对一个具体的技术方案而言, 所称汉字特定序位上的笔画究竟某个汉字的第几笔是确定的, 当然同一产品中也可以设置多个具体技术方案并存, 由用户根据需要进行选择所需的方案, 如一个方案取首笔, 另一个方案取首笔或者末笔等。该笔画或与汉字基本笔画建立对应关系的其他符号系列中对应于该笔画的符号均可以定义为相应汉字的编码。需要用到的汉字及其编码和区位码预先储存于有关的电子信息处理器中。如采用五笔划输入法输入汉字字串, 则汉字编码可以隐藏在该汉字的输入码中, 即电子信息处理器自动从该汉字的输入码中取出特定序位上的笔画作为其编码, 如: 取出某汉字五笔划输入码的头两码作为该汉字的编码。取出的汉字字串中汉字的编码组合形成相应数据项的检索码, 以及根据键盘输入的检索码完成检索功能由软件来实现。由于本发明技术方案设计精巧且对硬件

设备要求不高，因而可以应用本发明的电子器具绝非现有的数种。

本发明在检索时，对输入的检索码具有特殊的容错功能，使用户不至于在输入检索码时，因输错高位码位找不到所需的数据项而反复操作。其容错功能是电子信息处理器将数据库中各数据项的检索码的各码位与键盘输入的检索码的对应码位比较后，按具有相同码位的数目的多少，先多后少地用一移项功能键顺序调出相应各数据项。如：有检索码为从“000”到“999”的数据库，当输入检索码“111”时，首先调出检索码为“111”的数据项，然后每输入一次移项键，将顺序逐一调出检索码分别如下的各数据项：112, 113, …… 119, 110, 121, 131, 141, …… 191, 101, 211, 311, …… 911, 011, 122, 123, …… 继续输入移项键时，没有相同码位的数据项可按输入顺序或码位的高低顺序调出。又如：某数据库中的数据项检索码分别为：111111, 121212, 123456, 131313, 141414, 151515, 161616, 223456，在检索时因判别或输入错误，本想检索检索码为：“223456”的数据项，但实际输入为“123456”，按现有排序方法则按以下顺序先后显示各数据项：123456, 131313, 141414, 151515, 161616, 223456, 111111, 121212，也就是需用五或六次移项功能键才能调出正确的数据项“223456”；如果按本发明所述的容错功能，则可按以下顺序先后显示各数据项：123456, 223456, 121212, 141414, 161616, 111111, 131313, 151515，显然，只需一或二次移项功能键即可调出所需数据项“223456”。

上述两种技术方案也可将汉字中的若干偏旁部首独立出来与分类得到的汉字基本笔画并列作为字元对汉字进行编码：如：“气”，

“木”，“彳”，“才”，“口”，“一”，“丨”，“丿”，“丶”，“乙”，如取每个字的头尾字元为该字的编码时“树”字的编码为“木\”；“周”字的编码为“丿口”；“吕”字的编码为“口口”。

本发明汉字笔画的基本分类可有以下几种：

1、汉字的基本笔画可分为“横”，“竖”，“撇”，“捺”；分别用以下符号表示：“一”，“丨”，“丿”，“丶”；其中“点”归在“捺”内，所有“折”笔根据其第一折弯的前段或最后一折弯的后段分别归入“一”，“丨”，“丿”，“丶”。

2、汉字的基本笔画可分为“横”，“竖”，“撇”，“点”，“捺”；另用以下符号表示“一”，“丨”，“丿”，“丶”，“丶”，“丶”；所有“折”笔根据其第一折弯的前段或最后一折弯的后段分别归入“一”，“丨”，“丿”，“丶”，“丶”，“丶”。

3、汉字的基本笔画分为：“横”，“竖”，“撇”，“点”，“折”；“一”，“丨”，“丿”，“丶”，“乙”；“捺”归在“点”内。

4、汉字的基本笔画分为：“横”，“竖”，“撇”，“点”，“捺”，“折”；分别用以下符号表示：“一”，“丨”，“丿”，“丶”，“丶”，“乙”。

5、汉字的基本笔画分为：“横”，“竖”，“撇”，“捺”，“顺折”，“反折”分别用以下符号表示：“一”，“丨”，“丿”，“丶”，“ㄣ”，“ㄤ”；其中“点”归在“捺”内，收笔时折向为顺时针方向的“顺折”，反之为“反折”。

6、汉字的基本笔画可分为“横”，“竖”，“撇”，“点”，“捺”，“顺折”，“反折”；分别用以下符号表示“一”，“丨”，“丿”，“丶”，“丶”，“ㄣ”，“ㄤ”；其中收笔时折向为顺时针方向的为“顺

折”，反之为“反折”。

7、汉字的基本笔画可分为“横”，“竖”，“撇”，“捺”，“单折”，“多折”，分别用以下符号表示“一”，“丨”，“丿”，“㇏”，“㇚”，“乙”；其中其中“点”归在“捺”内，“单折”为只有一个弯的折笔，“多折”为具有二个或二个以上转弯的折笔。

8、汉字的基本笔画可分为“横”，“竖”，“撇”，“点”，“捺”；“单折”，“多折”；分别用以下符号表示“一”，“丨”，“丿”，“、”，“丶”；“㇚”，“乙”；其中“单折”为只有一个弯的折笔，“多折”为具有二个或二个以上转弯的折笔。

以上分类方法使用起来各有优缺点，但是可以肯定将“点”和“捺”分开的分类方法不是本发明的优先方案，原因在于汉字中“点”和“捺”有时很难区分。

本发明汉字基本笔画在键盘位上的分布不受本发明技术方案的限制，完全可以根据具体硬件的特点进行定义，也就是说汉字笔画在键盘位上的定义布是可以变化的。

以下是其中两种键盘定义方案：

1、汉字的每个基本笔画用单个按键定义。

2、汉字的每个基本笔画用二键次或二键次以上的按键组结合笔画的形态来定义，以下是其中的一些实例：

(1) 汉字的基本笔画分为“横”，“竖”，“撇”，“捺”；分别用以下符号表示：“一”，“丨”，“丿”，“㇏”；“折”笔根据其第一个转弯的前段分别归入“横”，“竖”，“撇”，“捺”各笔中；键盘上同一行中不同的两个按各输入一次定义为“横”的模拟键组，同一列中不同的两个键各输入一次定义为“竖”的模拟键组，不同行不同列其连线

方向自上而下向左边的两个键各输入一次定义为“撇”的模拟键组，不同行不同列其连线方向自上而下指向右边的两个按键各输入一次定义“捺”的模拟键组。

(2) 汉字的基本笔画分为“横”，“竖”，“撇”，“捺”；分别用以下符号表示：“-”“|”“丿”，“\”，“折”笔根据其最后一个折弯的后段分别归入“横”，“竖”，“撇”，“捺”各笔中；键盘上同一行中不同的两键各输入一次定义为“横”的模拟键组，同一列中不同的两个键各输入一次定义“竖”的模拟键组，不同行不同列其连线方向自上而下向左边的两个键各输入一次定义为“撇”的模拟键组，不在同行不同列其连线方向自上而下指向右边的两个按键各输入一次定义为“捺”的模拟键组。

(3) 汉字的基本笔画分别“横”，“竖”，“撇”，“捺”，“点”；分别用以下符号表示：“-”“|”“丿”，“\”，“、”，“折”笔根据其第一个折弯的前段分别归入“横”，“竖”，“撇”，“捺”，“点”，中；键盘上同一行中不同的两键各输入一次定义为“横”的模拟键组，同一列中不同的两个键各输入一次定义为“竖”的模拟键组，不在同一行也不在同一列其连线方向自上而下向左边的两个键各输入一次定义为“撇”的模拟键组，不在同行也不在同一列其连线方向自上而下指向右边的两个按键各输入一次定义为“捺”的模拟键组，任一按键连续输入两次定义为“点”的模拟键组。

(4) 汉字的基本笔画分别“横”，“竖”，“撇”，“捺”，“点”；分别用以下符号表示：“-”“|”“丿”，“\”，“、”，“折”笔根据其最后一个折弯的后段分别归入“横”，“竖”，“撇”，

“捺”，“点”，中；键盘上同一行中不同的两键各输入一次定义为“横”的模拟键组，同一列中不同的两个键各输入一次定义为“竖”的模拟键组，不在同一行也不在同一列其连线方向自上而下向左边的两个键名输入一次定义为“撇”的模拟键组，不在同行也不在同一列其连线方向自上而下指向右边的两个按键各输入一次定义为“捺”的模拟键组，任一按键连续输入两次定义为“点”的模拟键组。

(5) 汉字的基本笔画分为“横”，“竖”，“撇”，“捺”，“折”；分别用以下符号表示：“-” “|” “丿”，“\”，“ㄣ”，“点”归在“捺”内；用键盘中位于“口”形四个角上的键(1)键(2)键(3)键(4)，(图4)来定义各笔画的模拟键组；键(1)和键(2)各输入一次组成“横”的模拟键组，键(2)和键(4)各输入一次组成“竖”的模拟键组，键(2)和键(3)各输入一次组成“撇”的模拟键组，由键(1)和键(4)或由键(1)、键(3)和键(4)各输入一次组成“捺”的模拟键组，由键(1)连续输入两次组成“折”的模拟键组。

(6) 汉字的基本笔画分为“横”，“竖”，“撇”，“捺”，“点”，“折”；分别用以下符号表示：“-” “|” “丿”，“\”，“、”，“ㄣ”；“点”归在“捺”内；用键盘中位于“十”形四端的键(5)、键(6)、键(7)、键(8) (图5)来定义各笔画的模拟键组，键(6)和键(8)各输入一次组成“横”的模拟键组，由键(5)和键(7)各输入一次组成“竖”的模拟键组，键(5)和键(6)各输入一次定义“撇”的模拟键组，由键(5)和键(8)各输入组成“捺”的模拟键组，键(5)连续输入二次组成“点”的模拟键组，键(6)、键(8)和键(7)各输入一次组成“折”的模拟键组。

(7) 汉字的基本笔画分为“横”，“竖”，“撇”，“捺”，“顺折”，“反折”；分别用以下符号表示：“-”“|”“J”，“\”，“┐”，“└”；“顺折”为第一个折向或最后一个折向为顺时针方向的折笔，反之为“反折”；“点”归在“捺”内；用键盘中位于“口”形四个角上的键(1)、键(2)、键(3)、键(4) (图4)来定义各笔画的模拟键组；键(1)和键(2)各输入一次组成“横”的模拟键组，键(2)和键(4)各输入一次组成“竖”的模拟键组，键(2)和键(3)各输入一次组成“撇”的模拟键组，键(1)和键(4)各输入一次组成“捺”的模拟键组，键(1)、键(2)和键(4)各一次组成“顺折”的模拟键组。键(1)、键(3)和键(4)各一次组成“反折”的模拟键组。

(8) 汉字的基本笔画分为“横”，“竖”，“撇”，“捺”，“顺折”，“反折”；分别用以下符号表示：“-”“|”“J”，“\”，“┐”，“└”；“顺折”为第一个折向或最后一个折向为顺时针方向的折笔，反之为“反折”；“点”归在“捺”内；用键盘中位于“十”形四端的键(5)、键(6)、键(7)、键(8) (图5)来定义各笔画的模拟键组，键(6)和键(8)各输入一次组成“横”的模拟键组，由键(5)和键(7)各输入一次组成“竖”的模拟键组，键(5)和键(6)各输入一次定义“撇”的模拟键组，由键(5)和键(8)各输入组成“捺”的模拟键组，键(5)、键(8)和键(7)各输入一次组成“顺折”的模拟键组，键(5)、键(6)和键(7)各输入一次组成“反折”的模拟键组。

对汉字进行编码时，最好根据以下三种方案选取某汉字的笔画作为编码：

1、对汉字进行编码时同时取出汉字的首笔笔画和末笔笔画，首笔笔画在前，末笔笔画在后，两者组合在一起定义为该汉字编码，对只有一个笔画的汉字可以只取该笔画或取二次该笔画作为汉字的编码。

2、对汉字进行编码时只取出汉字的首笔笔画，首笔笔画即定义为该汉字的编码，只有一个笔画的汉字只取该笔画作为汉字的编码。

3、对汉字进行编码时只取出汉字的末笔笔画，末笔笔画即定义为该汉字的编码，只有一个笔画的汉字只取该笔画作为汉字的编码。

4、对汉字进行编码时同时取出汉字的首笔笔画和第二笔笔画，首笔笔画在前，第二笔笔画在后，两者结合在一起定义为该汉字的编码，对只有一个笔画的汉字可以只取该笔画或取两次该笔画作为汉字的编码。

另外，根据作为数据库检索入口的汉字字符串来确定相应数据项的检索码也有多种方案，但大多是由于选取汉字字符串中不同数目的汉字而产生的。以下四种方案是最可行的：

1、电子信息处理器取出作为数据库检索入口而输入到该电子信息处理器的汉字字符串的每个汉字的编码，取出的汉字编码根据相应汉字在汉字字符串中的顺序组合形成相应数据项的检索码。

2、电子信息处理器也可以取出作为数据库检索入口而输入到该电子信息处理器的汉字字符串中的首位汉字和末位汉字的编码，取的汉字编码根据相应汉字在汉字字符串中的顺序组合形成相应数据项的检索码。

3、电子信息处理器取出作为数据检索入口而输入到该电子信息处理器的汉字字符串中的前两个汉字的编码，对只有一个汉字的字符串只取一个汉字编码，取出的汉字编码根据相应汉字在汉字字符串中的顺序组合形成相应数据项的检索码。

4、电子信息处理器取出作为数据库检索入口而输入到该电子信息处理器的汉字字符串中的前三个汉字的编码，对只有一或两个汉字字符串的，取全部汉字的编码，取出的汉字编码根据相应汉字在汉字字符串中的顺序组合形成相应数据项的检索码。

虽然本发明检索时的重码率很低，但少量重码现象仍不可避免，此类问题可由软件来解决，例如显示出重码的数据项，或进行提示，然后由操作者进行选择。也可以在资料输入过程中边输入边判别重码。发现重码时提示用户，由用户修改输入的字符串，如：已存入实施例2中以“陈大明”为字符串的数据资料，再要输入以“陈大明”为字符串的资料时，机器提示为重码，用户可改为“陈大明二”的字符串输入，以便区别。

本发明考虑到由于手提电话、传呼机、个人移动信息处理器等产品的汉字字库不可能将全部汉字都收藏，所以对字库中找不到的汉字，取本发明所述的编码表示该汉字，也就是说上述两方案中之A所述的汉字包括汉字基本笔画：“一”“丨”“丿”，“丶”，“㇏”，“㇏”，“乙”和汉字部首：“彳”，“木”，“讠”，“扌”，“口”，“火”，“土”，“禾”，“竹”，“文”，“灬”，“辶”等。

本发明对夹杂在汉字字符串中的英文字母等字符可以作无字符处理，也可以用单个按键集中只用一个编码处理，也可以将这些

字符分类后用若干特定编码定义，还可能对字符的笔画分类，取其
与汉字一致的编码规则。

本发明的编码检索方法，不受汉字偏旁部首的限制，与汉语
拼音无关，无需拆字，只需记忆汉字的几个基本笔画并能正确掌
握汉字的书写笔画顺序即随意操作。因而，本发明具有直观、便
捷、记忆量小等优点。

下面结合实施例对本发明作进一步说明。

以下实施例中汉字基本笔画分为：“横”，“竖”，“撇”，“
捺”，“折”，即“-”“|”“丿”，“\”，“乙”；汉字的编
码采用同时取该字的首笔画和末笔画方法，汉字字串相应数据项
的检索码取该字串所有汉字的编码组成。当然，选择的方案还
有很多，但各种方案都是类似的，本领域普通技术人员在本说明
书的基础上完全可以实施。故本说明书不再赘述。

实施例 1：

汉字字串“汽车”中，“汽”的编码是“\乙”，“车”的
编码是“-|”，字串“汽车”作为数据库检索入口时，相应数
据项的检索码就是“\乙-|”。检索时，只需从键盘输入检索
码“\乙-|”，即可检索到以汉字字串“汽车”作检索口的数
据项资料。

实施例 2：

汉字字串“陈大明”中，“陈”的编码是“乙\”，“大”的
编码是“-\”，“明”的编码是“|一”，则字串“陈大明”作数
据库检索入口时，相应数据项的检索码就是“乙\-\|一”。
检索时，只需从键盘输入检索码“乙\-\|一”，即可检索到

以汉字字串“陈大明”作检索口的数据项资料。

· 实施例 3:

汉字字串“欧阳天志”中，“欧”的编码是“一\”，“阳”的编码是“乙一”，“天”的编码是“一\”，“志”的编码是“一\”，则字串“欧阳天志”作数据库检索入口时，相应数据项的检索码就是“一\乙一一\一\”。检索时，只需从键盘输入检索码“一\乙一一\一\”，即可检索到以汉字字串“欧阳天志”作检索口的数据项资料。

· 实施例 4:

汉字字串“广东”中，“广”的编码是“\丿”，“东”的编码是“一\”，则字串“广东”，作数据库检索入口时，相应数据项的检索码就是“\丿一\”。检索时，只需从键盘输入检索码“\丿一\”即检索到以字串“广东”作检索入口的数据项资料。

附图 1 是汉字基本笔画定义在电话机、传真机键盘数字键上的示意图。

附图 2 是汉字基本笔画定义在电话机、传真机键盘专门笔画键上的示意图。

· 附图 3 是电话机、传真机键盘汉字基本笔画部分定义在数字键或功能键上，部分定义在专门的笔画键上的示意图。

附图 4 是汉字基本笔画根据其形态定义在键盘“口”形四个角上的键(1)、键(2)、键(3)、键(4)上的示意图。

附图 5 是汉字基本笔画根据其形态定义在键盘“十”形四端上的键(5)、键(6)、键(7)、键(8)的示意图。

说 明 书 附 图

1	2	3
4 J	5	6 \
7	8 乙	9
*	0	#

图 1

—	1	2	3
	4	5	6
J	7	8	9
\	*	0	#
乙			

图 2

93-12-29

1	2	3
4	5	6
7	8	9
*	0	乙#
/		\

图3

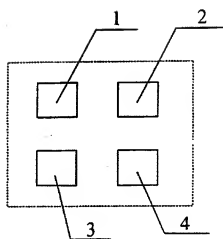


图 4

98-12-25

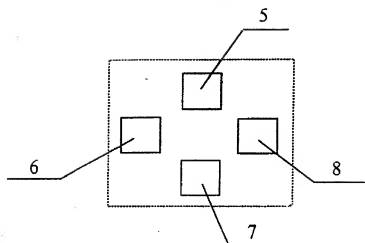


图 5